

Technical Report Number 08-04

**Examining Item Functioning of Math Screening Measures for Grades
1-8 Students**

Kimy Liu

Leanne R. Ketterlin-Geller

Paul Yovanoff

Gerald Tindal

University of Oregon

Behavioral Research and Teaching



behavioral research & teaching

Published by

Behavioral Research and Teaching
University of Oregon • 175 Education
5262 University of Oregon • Eugene, OR 97403-5262
Phone: 541-346-3535 • Fax: 541-346-5689
<http://brt.uoregon.edu>

Acknowledgments: We would like to thank the following individuals for their assistance in completing this research and preparing the report:

- Jenelle Braun-Monegan
- Drew Braun

Copyright © 2008. Behavioral Research and Teaching. All rights reserved. This publication, or parts thereof, may not be used or reproduced in any manner without written permission.

The University of Oregon is committed to the policy that all persons shall have equal access to its programs, facilities, and employment without regard to race, color, creed, religion, national origin, sex, age, marital status, disability, public assistance status, veteran status, or sexual orientation. This document is available in alternative formats upon request.

Abstract

BRT Math Screening Measures focus on students' mathematics performance in grade-level standards for students in grades 1-8. A total of 24 test forms are available with three test forms per grade corresponding to fall, winter, and spring testing periods. Each form contains computation problems and application problems. BRT Math Screening Measures were administered to 6,500 students during the 2006-07 school year. The Rasch Model in Item Response Theory (IRT) is employed to estimate the item difficulties and fit statistics of the test items. We describe this process and then present results of item difficulty and item functioning.

Introduction

Uses of tests and assessments have been key features of educational reform for over 50 years (Linn, 2000). Since the 1990's, the emphasis of assessment has shifted from documenting performance on norm-referenced tests to performance in standard-based accountability systems (Elmore, Abelmann, & Fuhrman, 1996; Linn, 2000). One of the ways to evaluate students' progress toward grade-level standards is to administer standardized general outcome measures (GOMs) (Shapiro, Keller, Lutz, Santoro, & Hintze, 2006). General outcome measures (GOMs) are designed for teachers to guide instructional design and delivery (Deno, 2003; Fuchs & Deno, 1991). Although GOMs have been used to evaluate students' academic performance for more than 30 years, most studies are in reading, particularly within the context of oral reading fluency (Buck & Torgeson, 2003; Crawford, Tindal, & Stieber, 2001; Good, Simmons, & Kame'enui, 2001). Few studies have reported the outcomes relationship between GOMs and statewide assessments in math (Shapiro et al., 2006). Helwig, Anderson and Tindals' (2002) study on the use of computerized adaptive assessment is an exception. In this study, students were provided with 48 math probes including problem solving and computation tasks. Results indicated that the probes predicted which students who would meet the state math standards with 87% accuracy (Helwig, Anderson, & Tindal, 2002). We have developed the BRT Math Screening Measures in a similar manner as a form of General Outcomes Measures for teachers to use in screening students who might be at risk for failure to meet expectations on the state content standards.

The purpose of this study is to describe how we used Item Response Theory (IRT) to examine the functionality of test items in the BRT Math Screening Measures. This technical report documents the procedures of estimating item difficulty and identifying problematic items that yielded unexpected response patterns. The test items that function appropriately can then

reliably generate expected patterns in which students with high skills not only answer more difficult questions correctly but they also answer more questions correctly. In contrast, students with low skills perform well only on easy items.

Methods

BRT Math Screening Measures align with the Oregon Mathematics Curriculum Standards and are designed to assess students' skills, knowledge and ability in the five domains of mathematics. In this study, these measures were administered to students in grades 1-8 in two local school districts in the Pacific Northwest Region in the fall, winter, and spring during the school year of 2006-2007.

Setting and Subjects

BRT Math Screening Measures were administered to students during their regular math instruction. Approximately 6,500 students participated in the study with the number varying widely across grades, ranging from 1,501 students in Grade 5 to 401 students in Grade 7 (see Table 1 for sample sizes per grade). Within each grade, the sample population fluctuated slightly between the fall, winter, and spring measures, possibly due to students' absences during the scheduled test weeks or mobility during the school year. No demographic data were available.

Measurement/Instrument Development

BRT Math Screening Measures contains two sections: computation problems and application problems. The computation questions assess student's procedural knowledge of algorithms and arithmetic. The application problems assess student's ability to apply mathematical concepts and algorithms to situated problems. Grade-level appropriate math vocabulary words (such as mean, mode and range) are included in the BRT Math Screening

Measures as specified by the content standards. With the exception of Grade 1 computation problems, that require students to show their work, all BRT Math Screening Measures utilize a multiple-choice format with four possible choices. Students are not penalized for incorrect guessing.

The BRT Math Screening Measures were administered in a paper-pencil format. The students answered questions on a separate answer sheet. The numbers of questions varied across grades. Students were not permitted to use calculators on computation problems, however, calculators were permitted on the application problems. Most students finished the assessment within a 45-minute session. If a student did not finish the test in one math session, they were asked to continue in the following session.

Data Preparation and Analysis

This technical report documents the procedure of data analysis, using Item Response Theory (IRT) and reports the results of statistical fit analyses of students' responses using WINSTEPS (Linacre, 2006) to analyze students' response patterns in all 24 tests (Grades 1-8, three tests per grade) and item difficulty of each test item (1PL). The procedure of data preparation and inspection are described in the following sections.

Data Preparation

The data procedure for the math problem preparation was as follows:

1. A combined data file for each of 24 measures was created in the Excel® format. The data files contained students' responses in computation problems and application problems.
2. The combined Excel® files were read into SPSS data files and then into the WINSTEP program.

3. A control file embedded with the answer keys was created for each of the 24 measures to allow automatic scoring of items in the WINSTEPS software.
4. After a number of iterations, the WINSTEPS software produced the fit statistic output files that indicated the estimated item difficulty for each test item.

Data Inspection

We examined the item characteristics of each item from the 24 fit statistic output files (Tables 2-25). For the test items to function appropriately, the students' response patterns must indicate that students who are more skillful not only answer more difficult questions correctly but also answer more questions correctly. Moreover, the response patterns need to fit what is expected or predicted in the proposed model, as evidenced by a reasonable number of students' scores falling within .5 score points of their expected values. To examine the functionality of a test item, special attention is directed toward the outfit mean square fit statistics of that item, which is more sensitive to unexpected behaviors by persons on items far from the person's ability levels. According the WINSTEPS Manual, items that have outfit square means between .5 and 1.5 are considered productive items because examinees' response patterns are within reasonable expectation (Linacre, 2006). The items with outfit mean squares outside this range are considered non-productive items. These non-productive items warrant further inspection.

Follow-up inspection entailed examining whether the answer keys for non-productive items was correct by conducting a distractor analysis. WINSTEPS output files reported the number of students choosing each multiple-choice option and the average ability level of students who choose each option. On average, the students who chose correct answers should have higher average ability level than the students who chose incorrect distractors. When the average ability level of students who chose the correct answer was lower than the average of

students who chose distractors, WINSTEPS placed an asterisk to signify an unexpected pattern in the distractor analysis. Such an unexpected response pattern prompted inspecting the correctness of the answer keys. If a test item had an incorrect answer key, we revised the answer key in the WINSTEP control file and re-ran the WINSTEP program. The revised output files for that particular measure has been reported following the original output files. For fall Grade 1 computation problems, distractor analysis was not available because they were constructed-response items.

Results

Most items in the BRT Math Screening Measures were deemed productive items, as evidenced by their outfit mean square fit statistics within the range of .5 and 1.5 (Tables 2-25). Of the approximately 1,000 test items, only 45 test items were identified as problematic. Three types of problematic items were identified:

1. Nine items had incorrect answer keys;
2. Six items were recommended for deletion or revision because outfit mean square exceeded 2.0 (Linacre, 2006).
3. Thirty items were kept in the test bank because the items yielded useful information and outfit mean square was within the range of 1.5 and 2.0.

All problematic items were identified and categorized in the Table 26. The items with incorrect keys were re-classified as productive items once correct answer keys were input for data analysis (e.g. see Table 5B).

To determine whether the measures were comparable in item difficulty within a grade-level, the author examined the range and distribution of measures (item difficulty) within each of the grade levels and across the fall, winter, and spring tests. In the output file of each measure,

the percentage of data points within .5 score points of their expected values has been reported in the table under the *Obs. Match*. The percentage of data points within .5 score point that were predicted to be within .5 score points of their expected value has been reported under the column *Exp. Match*. High percentage of these two numbers indicated the data fitted with the proposed Rasch model.

Finally, students' scale scores were obtained in the WINSTEP output files. The comparison of students' scale scores on BRT Math Screening Measures within the grade level indicated students' scores increased over time. Such comparison was valid and meaningful because (a) the overwhelming majority of students took all three tests over the course of school year and (b) linking items were embedded across the measures within the grade.

Discussion

BRT Math Screening Measures were administered to approximately 6,500 students in Grades 1-8 in the fall, winter, and spring of 2006-07. The number of items varied across grades, ranging from 20 items in Grade 1 fall measures to 50 items in the Grades 6-8 measures. Almost all of the nearly 1,000 test items on the BRT Math Screening Measures functioned well and yielded predictable response patterns in which the students who chose correct answers as a group were more skillful or had more math knowledge and ability than students who answered the question incorrectly. Of 45 possible problematic items, only 6 items had the outfit mean squares exceeding 2.0 and were deleted from the test bank. Had we deleted all problematic items without close analysis, 39 test items would have been deleted.

References

Buck, J., & Torgeson, J. (2003). *The relationship between performance on a measure of oral reading fluency and performance on the Florida comprehensive assessment test (technical report 1)*. Tallahassee, FL.: Center for Reading Research.

Crawford, L., Tindal, G., & Stieber, S. (2001). Using oral reading rate to predict student performance on statewide achievement tests. *Educational Assessment*, 7, 303-323.

Deno, S. D. (2003). Development in curriculum based measurement. *Journal of Special Education*, 37, 184-192.

Elmore, R. F., Abelmann, C. H., & Fuhrman, S. H. (1996). The new accountability in state education reform: From process to performance. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 65-98). Washington, DC: The Brookings Institute.

Fuchs, L. S., & Deno, S. (1991). Paradigmatic distinctions between instructional relevant measurement models. *Exceptional Children*, 57, 488-500.

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of continuum of fluency-based indicators of foundational reading skills for third grade high-stakes outcomes. *Scientific Studies of Reading*, 5(3), 257-288.

Helwig, R., Anderson, L., & Tindal, G. (2002). Using a concept-grounded, curriculum-based measure in mathematics to predict statewide test scores for middle school students with ID. *Journal of Special Education*, 36(2), 102-112.

Linacre, J. B. (2006). Winsteps 3.61.1. Rasch-model computer programs [computer software]. Chicago, IL: MESA.

Linn, R. L. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4-16.

Shapiro, E. S., Keller, M. A., Lutz, J. G., Santoro, L. E., & Hintze, J. M. (2006). Curriculum-based measures and performance on state assessment and standardized tests: Reading and math performance in Pennsylvania. *Journal of Psychoeducational Assessment*, 24(1), 19-35.

Table 1.
Number of Participants and Test Items in BRT Math GOMs.

Grade	Fall		Winter		Spring	
	Participants	Items	Participants	Items	Participants	Items
1	1262	20	1308	30	1379	30
2	1382	40	1389	45	1465	45
3	1393	50	1392	50	1447	50
4	908	50	1409	50	1398	50
5	1507	50	1467	50	1490	50
6	466	50	496	50	491	50
7	401	50	437	50	426	50
8	455	50	456	55	454	50

Tables 2-25

Note: The output files of BRT Math Screening Measures reported item number, estimated item difficulty (shown in the tables as Measures), number of students who attempted to answer the questions (shown in the tables as Count), and number of students who answered the questions correctly (shown in the tables as Score). The output reported (a) outfit mean square (shown in the table as OUTFIT MNSQ), and (b) outfit mean-square fit statistics t-standardized (shown in the table as OUTFIT MNSQ ZSTD) for each test items. The output also reported the percentage of observed data points which are within .5 score points of their expected values and the percentage of data points which are predicted to be within .5 score points of their expected values (shown in the table as OBSMATCH and EXPMATCH respectively).

Table 2.
Grade 1 Fall 2006 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	-3.67	1262	1143	4.76	8.16	90.7	92.8
2	-2.79	1262	1075	0.67	-1.73	90.4	90.6
3	-3.21	1262	1110	0.8	-0.84	90.8	91.7
4	-0.95	1262	861	1.75	4.65	79.6	84
5	-2.09	1262	1006	2.02	4.5	88.5	88.2
6	-2.51	1262	1049	0.63	-2.07	92.9	89.7
7	2.37	1262	306	8.5	9.91	83.1	84
8	-1.96	1262	992	0.87	-0.73	89.8	87.8
9	3.1	1262	211	1.71	3.25	89.9	87.8
10	2.31	1262	315	2.62	7.09	87.3	83.6
11	5.77	1262	36	1.08	0.38	97.1	97.1
12	-0.45	1262	784	1.07	0.6	82.9	82
13	0.68	1262	591	0.94	-0.43	84.7	78.9
14	-1.79	1262	972	0.65	-2.29	89.8	87.1
15	0.54	1262	616	1.13	1.04	82.1	79.3
16	3.03	1262	220	1.85	3.84	88.1	87.4
17	-1.25	1262	903	0.67	-2.52	88.2	85.2
18	2.6	1262	273	2.03	4.73	88.1	85.2
19	1.23	1262	491	0.97	-0.13	80.6	78.9
20	-0.96	1262	863	0.93	-0.5	86.8	84

Table 3.
Grade 2 Fall 2006 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	38.57	900	771	0.80	-1.69	86.80	86.10
2	48.96	812	581	0.89	-1.48	74.40	75.80
3	35.15	924	820	1.19	1.23	88.90	89.00
4	49.80	822	576	0.88	-1.82	76.30	75.00
5	37.65	849	736	0.85	-1.08	88.00	87.10
6	60.18	749	379	0.94	-1.32	72.60	69.70
7	46.28	891	676	0.84	-2.02	83.30	78.40
8	33.94	907	816	0.65	-2.46	89.90	90.10
9	44.58	904	707	1.07	0.78	82.40	80.00
10	46.57	854	643	0.82	-2.33	80.80	78.10
11	61.83	783	374	0.82	-3.95	76.10	69.40
12	54.60	774	477	0.83	-3.26	76.40	71.40
13	62.00	772	362	0.99	-0.27	70.50	69.60
14	64.60	743	312	1.20	3.60	65.40	70.60
15	63.91	698	301	0.97	-0.47	74.80	70.60
16	65.26	702	283	1.55	8.40	63.70	70.80
17	56.18	686	400	0.94	-1.09	71.60	70.40
18	67.11	658	237	1.47	6.43	68.40	72.30
19	44.76	803	627	0.85	-1.62	81.40	80.00
20	27.48	867	816	0.85	-0.64	94.00	94.10
21	37.21	905	787	1.12	0.90	87.50	87.30
22	48.46	842	613	0.85	-2.10	77.80	76.20
23	46.82	868	649	1.07	0.89	77.60	77.60
24	21.34	908	877	0.81	-0.62	96.60	96.60
25	51.37	891	597	0.84	-2.92	77.40	73.10
26	36.81	898	786	1.04	0.32	88.10	87.80
27	57.65	886	488	1.09	1.90	66.70	69.20
28	46.46	864	651	0.91	-1.17	78.50	77.90
29	48.75	897	643	1.12	1.76	72.00	75.50
30	62.59	888	399	1.28	5.86	64.50	69.40
31	49.20	904	638	1.18	2.58	72.90	75.00
32	33.82	893	805	0.62	-2.67	90.50	90.20
33	61.55	887	416	1.15	3.38	66.40	69.00
34	57.78	848	462	1.06	1.24	68.60	69.00
35	60.45	886	439	1.27	5.75	60.60	68.80
36	71.80	830	233	1.27	3.42	73.40	75.60
37	47.48	892	653	1.11	1.50	77.50	76.50
38	52.90	868	556	0.79	-4.07	78.00	71.70
39	48.72	892	637	0.95	-0.67	74.90	75.50
40	49.43	875	612	0.93	-1.11	73.80	74.60

Table 4.
Grade 3 Fall 2006 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	36.40	1348	1074	0.77	-3.12	81.60	80.60
2	35.97	1348	1082	0.92	-0.96	81.20	81.10
3	40.68	1348	986	0.75	-4.45	78.90	75.70
4	64.76	1348	373	1.21	3.50	76.30	76.20
5	54.07	1348	642	0.80	-6.12	77.60	68.80
6	52.45	1348	686	1.00	0.08	68.20	68.60
7	51.35	1348	716	0.84	-4.81	77.00	68.60
8	57.93	1348	539	0.82	-4.83	78.50	70.40
9	69.17	1348	283	1.32	3.97	82.60	80.90
10	50.32	1348	744	0.92	-2.09	71.40	68.70
11	45.65	1348	867	0.83	-3.84	73.40	71.10
12	45.46	1348	872	1.07	1.53	68.60	71.20
13	46.24	1348	852	0.83	-4.15	76.50	70.60
14	45.26	1348	877	0.72	-6.68	78.50	71.40
15	68.27	1348	300	1.28	3.69	82.10	79.90
16	41.39	1348	970	0.84	-2.98	76.70	75.00
17	39.30	1348	1016	0.75	-4.11	79.50	77.30
18	64.71	1348	374	1.24	3.99	76.60	76.20
19	55.44	1348	605	1.21	5.45	62.00	69.20
20	58.66	1348	520	0.97	-0.82	72.70	70.90
21	33.11	1348	1131	1.55	5.03	84.30	84.40
22	59.99	1348	486	1.26	5.60	66.30	71.80
23	61.45	1348	450	0.99	-0.11	75.00	73.00
24	44.58	1348	894	0.90	-2.23	74.60	71.90
25	56.19	1348	585	1.47	9.90	58.10	69.50
26	47.36	1348	823	1.01	0.16	67.60	69.90
27	55.25	1348	610	1.05	1.28	67.10	69.10
28	62.87	1348	416	1.04	0.78	75.60	74.30
29	44.58	1348	894	1.42	7.66	64.90	71.90
30	63.56	1348	400	1.13	2.36	76.00	75.00
31	47.05	1348	831	0.83	-4.24	73.60	70.10
32	66.58	1348	334	1.46	6.40	75.30	78.10
33	41.22	1348	974	0.92	-1.43	77.40	75.20
34	53.89	1348	647	1.09	2.54	65.40	68.80
35	31.54	1348	1155	0.92	-0.75	86.60	86.00
36	47.20	1348	827	1.14	3.32	67.70	70.00
37	51.39	1348	715	1.06	1.63	68.00	68.60
38	54.18	1348	639	0.96	-1.24	69.80	68.80
39	30.11	1348	1175	0.75	-2.48	88.10	87.40
40	47.20	1348	827	0.84	-4.11	75.10	70.00
41	66.24	1348	341	1.17	2.69	74.90	77.80
42	40.32	1348	994	0.92	-1.34	78.00	76.10
43	51.98	1348	699	1.07	1.93	69.10	68.50
44	48.87	1348	783	1.19	4.60	65.50	69.20
45	54.70	1348	625	1.26	6.70	62.40	69.00
46	30.55	1348	1169	0.98	-0.14	87.80	87.00
47	55.89	1348	593	1.01	0.30	68.70	69.40
48	30.84	1348	1165	0.78	-2.21	87.40	86.70
49	53.48	1348	658	1.01	0.37	68.00	68.70
50	44.38	1348	899	0.91	-1.91	73.30	72.10

Table 5A.
Grade 4 Fall 2006 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	24.56	882	802	1.26	1.59	90.5	91.6
2	42.04	882	612	0.94	-1.08	74.4	73.7
3	46.37	882	540	0.77	-5.32	74.4	69.3
4	47.06	882	528	0.94	-1.18	69.2	68.8
5	33.51	882	726	0.78	-2.43	84.1	83.6
6	53.55	882	410	0.97	-0.66	68.9	67.1
7	66.78	882	194	1	-0.01	78	79.2
8	45.45	882	556	0.79	-4.64	75.7	70.1
9	59.88	882	298	0.98	-0.33	76.4	71
10	36.24	882	694	1.11	1.37	81.2	80.6
11	47.06	882	528	0.8	-4.47	73.7	68.8
12	55.31	882	378	1.09	1.64	66.7	67.6
13	56.98	882	348	0.99	-0.13	71.4	68.6
14	29.18	882	768	0.68	-2.86	87.8	88
15	45.56	882	554	0.94	-1.14	71.4	70
16	62.48	882	256	1.23	2.79	72.8	73.9
17	44.14	882	578	0.95	-0.95	71.7	71.4
18	41.91	882	614	0.81	-3.38	74.1	73.9
19	52.24	882	434	0.88	-2.76	71.7	67
20	44.74	882	568	0.98	-0.33	71.2	70.8
21*	96.02	882	16	3.27	4.44	98.2	98.2
22	57.43	882	340	1.04	0.64	69.6	68.9
23	33.69	882	724	0.71	-3.38	84.4	83.4
24	39.79	882	646	0.78	-3.55	79.4	76.4
25	67.09	882	190	1.34	3.34	79.8	79.5
26	41.00	882	628	0.75	-4.44	76.9	74.9
27	68.52	882	172	1.05	0.48	83	81.3
28	50.93	882	458	1.03	0.74	66.2	67.2
29	49.62	882	482	1.07	1.45	63.9	67.5
30	57.55	882	338	1.01	0.21	72.1	69
31	61.34	882	274	1.26	3.33	73.7	72.6
32*	60.36	882	290	1.71	8.71	56	71.5
33	44.14	882	578	1.01	0.15	73.5	71.4
34	54.21	882	398	1.09	1.82	64.6	67.3
35	63.66	882	238	1.76	7.83	66.7	75.3
36	54.10	882	400	1.08	1.57	65.3	67.2
37	49.18	882	490	0.94	-1.37	67.6	67.7
38	38.95	882	658	0.76	-3.68	80.3	77.4
39	46.14	882	544	0.9	-2.15	71.2	69.5
40	45.91	882	548	1.14	2.71	67.1	69.7
41	66.33	882	200	1.28	2.88	77.8	78.6
42	54.97	882	384	1.03	0.56	66	67.5
43	43.17	882	594	0.94	-1.06	75.7	72.5
44	48.18	882	508	0.94	-1.24	70.5	68.1
45	31.78	882	744	0.66	-3.6	86.8	85.4
46	54.31	882	396	0.95	-1.01	69.8	67.3
47	69.53	882	160	1.43	3.67	81	82.4
48	38.80	882	660	0.81	-2.93	80.5	77.6
49	47.96	882	512	1.17	3.43	66	68.3
50	30.31	882	758	0.89	-0.96	87.5	86.9

* indicates the wrong answer key has been used in the initial analysis. The correct revision was shown in the next page.

Note: After the correction (changing the answer key from D to A for item 21 and changing the answer key from A to B in Item 32) the revised data was shown in the following table.

Table 5B.
Grade 4 Fall 2006 Data After Correction.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	25.87	882	802	1.31	1.78	90.5	91.6
2	43.45	882	612	0.97	-0.54	73.7	74
3	47.82	882	540	0.78	-4.73	74.1	69.6
4	48.51	882	528	0.98	-0.44	68	69.1
5	34.87	882	726	0.8	-2.09	83.9	83.7
6	55.07	882	410	0.97	-0.66	69.8	67.5
7	68.46	882	194	1.03	0.32	78	79.4
8	46.89	882	556	0.8	-4.04	76.2	70.4
9	61.47	882	298	0.99	-0.09	75.5	71.5
10	37.61	882	694	1.14	1.59	81.4	80.7
11	48.51	882	528	0.8	-4.25	73.9	69.1
12	56.84	882	378	1.12	2.21	67.6	68.1
13	58.54	882	348	1	0.02	72.1	69.2
14	30.51	882	768	0.71	-2.52	88	88
15	47.01	882	554	0.97	-0.52	71.7	70.3
16	64.1	882	256	1.24	2.88	73.2	74.3
17	45.57	882	578	0.96	-0.75	70.1	71.7
18	43.32	882	614	0.82	-3.06	74.8	74.1
19	53.75	882	434	0.89	-2.22	72.3	67.4
20	46.18	882	568	1	-0.02	71.7	71.1
21	34.49	882	730	0.94	-0.52	85.3	84.1
22	58.99	882	340	1.05	0.84	69.8	69.5
23	35.05	882	724	0.7	-3.3	84.6	83.5
24	41.19	882	646	0.79	-3.22	78.9	76.5
25	68.77	882	190	1.37	3.5	79.1	79.8
26	42.41	882	628	0.74	-4.33	78	75.1
27	70.22	882	172	1.03	0.3	83.9	81.5
28	52.43	882	458	1.06	1.26	65.8	67.5
29	51.1	882	482	1.11	2.15	62.6	67.8
30	59.11	882	338	1.03	0.54	74.1	69.6
31	62.95	882	274	1.29	3.58	72.6	73
32	50.54	882	492	1.05	1.07	63.5	68
33	45.57	882	578	1.02	0.38	74.6	71.7
34	55.73	882	398	1.11	2.01	62.1	67.7
35	65.3	882	238	1.84	8.37	66	75.7
36	55.62	882	400	1.11	2.14	64.6	67.7
37	50.65	882	490	0.96	-0.89	66.9	68
38	40.34	882	658	0.77	-3.43	80.5	77.5
39	47.59	882	544	0.91	-1.73	71	69.8
40	47.36	882	548	1.17	3.22	65.5	70
41	68.01	882	200	1.26	2.68	77.8	78.9
42	56.51	882	384	1.07	1.29	66	68
43	44.59	882	594	0.93	-1.21	75.7	72.7
44	49.65	882	508	0.94	-1.16	71	68.4
45	33.13	882	744	0.67	-3.28	87.3	85.5
46	55.84	882	396	0.98	-0.42	68.3	67.7
47	71.25	882	160	1.49	4.07	81	82.7
48	40.2	882	660	0.8	-2.9	81.2	77.7
49	49.42	882	512	1.21	3.99	66.9	68.6
50	31.65	882	758	0.89	-0.92	87.3	87

Table 6.
Grade 5 Fall 2006 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	27.9	1461	1320	0.93	-0.57	90.4	90.4
2	47.24	1461	916	0.83	-4.82	73.6	69.1
3	45.1	1461	977	0.89	-2.72	72.8	71
4	39.62	1461	1117	0.94	-1.05	78.2	77.4
5	38.37	1461	1145	1.04	0.57	78.2	79.1
6	47.18	1461	918	0.78	-6.41	74.7	69.1
7	57.22	1461	614	0.99	-0.17	70.3	68.2
8	60.81	1461	510	1.19	4.88	70.8	70.9
9	63.86	1461	428	1.01	0.3	77.7	74.1
10	51.68	1461	783	0.99	-0.41	68.9	66.9
11	63.97	1461	425	1.18	3.73	76.1	74.2
12	45.46	1461	967	0.83	-4.22	74.5	70.6
13	48.94	1461	866	0.82	-5.66	75.4	67.9
14	65.54	1461	386	1.44	7.82	70.4	76.1
15	41.21	1461	1079	0.96	-0.79	75.7	75.4
16	51.84	1461	778	0.98	-0.77	68.4	66.8
17	35.11	1461	1211	0.8	-2.78	82.8	83.1
18	38.74	1461	1137	0.72	-4.81	80.2	78.6
19	63.2	1461	445	1.04	0.92	76.7	73.3
20	40.39	1461	1099	0.84	-2.86	77.2	76.4
21	41.25	1461	1078	1.01	0.14	76	75.3
22	41.62	1461	1069	0.92	-1.52	75.4	74.8
23	45.46	1461	967	1.25	5.65	65.3	70.6
24	60.56	1461	517	1.14	3.79	68.1	70.6
25	32.13	1461	1262	0.74	-3.15	86.9	86.5
26	55.53	1461	665	1.26	8.1	57	67.3
27	43.59	1461	1018	1.03	0.59	71.8	72.5
28	74.07	1461	214	1.06	0.79	86.4	85.9
29	60.27	1461	525	1.17	4.53	67	70.4
30	68.56	1461	317	1.37	5.48	80.3	79.8
31	54.75	1461	689	0.96	-1.22	67.9	67.1
32	46.55	1461	936	0.84	-4.35	73.9	69.6
33	58.68	1461	571	1	0.08	68.7	69.1
34	48.97	1461	865	1.12	3.45	66.1	67.9
35	52.4	1461	761	1.19	5.98	60.3	66.8
36	68.47	1461	319	0.98	-0.29	80.4	79.7
37	41.25	1461	1078	0.92	-1.45	76.5	75.3
38	65.91	1461	377	1.54	9.16	72.8	76.5
39	56.62	1461	632	1.07	2.32	64.8	67.8
40	29.34	1461	1302	0.88	-1.12	89.3	89.2
41	44.59	1461	991	0.94	-1.29	73.3	71.5
42	64.29	1461	417	1.35	6.82	71	74.6
43	55.34	1461	671	0.9	-3.52	73	67.3
44	48	1461	894	0.82	-5.36	74.1	68.5
45	52.53	1461	757	0.99	-0.19	67.1	66.8
46	47.72	1461	902	0.87	-3.67	73.6	68.7
47	43.89	1461	1010	0.84	-3.57	75.6	72.2
48	43.93	1461	1009	1.16	3.24	71.1	72.1
49	40.6	1461	1094	1.05	0.92	76.7	76.1
50	39.75	1461	1114	1.18	2.9	76.5	77.3

Table 7A.
Grade 6 Fall 2006 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	39.39	462	377	1.32	2.01	81.6	82.4
2	66.42	462	150	1.1	1.22	72.1	73.4
3	63.83	462	172	1.19	2.52	68	71.4
4	38.01	462	385	0.74	-1.8	87	83.9
5	59.4	462	212	0.77	-4.12	81.2	69
6	44.04	462	346	0.64	-3.71	78.8	77.1
7	54.77	462	255	0.67	-5.84	82	68.9
8	29.01	462	424	0.88	-0.42	92.2	91.9
9	54.77	462	255	0.79	-3.62	76.4	68.9
10	51.14	462	288	0.84	-2.21	71	70.4
11	40.21	462	372	0.65	-2.84	82.3	81.5
12	47.88	462	316	0.86	-1.62	75.8	73
13	42.2	462	359	0.68	-2.8	82	79.2
14	64.41	462	167	1.42	5.17	63.2	71.8
15	52.92	462	272	0.76	-3.8	76	69.5
16	46.27	462	329	0.88	-1.2	78.1	74.6
17	32.6	462	411	0.91	-0.39	89.2	89.1
18	55.2	462	251	1.12	1.83	64.7	68.8
19	53.25	462	269	1	-0.02	69.7	69.4
20	33.79	462	406	0.89	-0.54	87.7	88.1
21	41.6	462	363	0.99	0	81.2	79.9
22	48.48	462	311	0.9	-1.17	76	72.5
23	68.81	462	131	0.89	-1.15	79.9	75.8
24	52.92	462	272	1.14	1.94	63	69.5
25	37.28	462	389	0.92	-0.43	86.8	84.6
26	48.72	462	309	1.22	2.47	68.2	72.3
27	61.47	462	193	1.09	1.41	73.4	69.8
28	66.54	462	149	1.7	7.1	60.2	73.5
29	54.99	462	253	1.06	0.93	66.5	68.8
30	56.49	462	239	1.34	5.1	56.9	68.6
31	40.21	462	372	0.99	-0.05	82.3	81.5
32*	92.21	462	23	5.21	6.76	94.8	95.1
33	48.12	462	314	1.08	0.97	68.8	72.8
34	41.75	462	362	1.14	1.07	78.8	79.7
35	48.48	462	311	1	-0.01	71.2	72.5
36	46.4	462	328	1.09	0.88	74	74.5
37	74.75	462	90	1.14	1.05	82.3	82
38	51.14	462	288	1.07	0.92	67.1	70.4
39	50.69	462	292	0.93	-0.88	74	70.7
40	42.05	462	360	0.88	-0.92	81.8	79.3
41	38.54	462	382	1.04	0.33	84.6	83.3
42	49.42	462	303	0.96	-0.46	72.1	71.7
43	32.6	462	411	0.65	-1.84	89.2	89.1
44	57.57	462	229	1.02	0.39	69	68.6
45	60.7	462	200	1.22	3.27	66	69.5
46	49.77	462	300	1.13	1.59	69.9	71.4
47	45.63	462	334	0.87	-1.27	76.2	75.3
48	53.25	462	269	1.02	0.3	69.3	69.4
49	39.72	462	375	0.77	-1.68	81.6	82
50	30.21	462	420	0.65	-1.65	91.3	91

* Question 32 has an incorrect answer key. The correct answer should be B, not A. After the correction, the question becomes a constructive item.

Table 7B.
Grade 6 Fall 2006 Data After Correction.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	40.11	461	376	1.35	2.17	81.6	82.4
2	67.29	461	149	1.11	1.36	72.7	73.7
3	64.68	461	171	1.21	2.79	68.1	71.6
4	38.72	461	384	0.75	-1.7	86.3	83.8
5	60.22	461	211	0.77	-4.06	81.3	69.2
6	44.77	461	345	0.63	-3.76	78.5	77
7	55.56	461	254	0.67	-5.9	82.2	69
8	29.71	461	423	0.9	-0.32	92.2	91.8
9	55.56	461	254	0.79	-3.54	76.6	69
10	51.91	461	287	0.85	-2.13	71.6	70.5
11	40.93	461	371	0.65	-2.82	82.4	81.4
12	48.63	461	315	0.84	-1.88	76.8	73.1
13	42.93	461	358	0.69	-2.73	83.5	79.1
14	65.26	461	166	1.45	5.41	61.8	72
15	53.7	461	271	0.75	-3.94	76.4	69.6
16	47.01	461	328	0.89	-1.1	77.4	74.6
17	33.31	461	410	0.9	-0.41	89.2	89.1
18	55.99	461	250	1.13	1.95	64.4	68.9
19	54.03	461	268	1.01	0.11	69.8	69.5
20	34.49	461	405	0.88	-0.55	87.6	88
21	42.33	461	362	1	0.07	81.1	79.8
22	49.23	461	310	0.91	-1.03	76.1	72.5
23	69.7	461	130	0.91	-0.97	80.9	76
24	53.7	461	271	1.15	2.11	62.5	69.6
25	37.99	461	388	0.94	-0.33	87	84.6
26	49.47	461	308	1.25	2.74	69.2	72.3
27	62.3	461	192	1.11	1.61	72.5	70
28	67.42	461	148	1.74	7.33	60.3	73.8
29	55.78	461	252	1.07	1.18	65.3	69
30	57.29	461	238	1.36	5.35	55.7	68.8
31	40.93	461	371	1.02	0.17	82.4	81.4
32	54.58	461	263	0.93	-1.1	72.2	69.3
33	48.87	461	313	1.1	1.11	69	72.9
34	42.48	461	361	1.16	1.22	78.7	79.7
35	49.23	461	310	1.02	0.3	71.4	72.5
36	47.14	461	327	1.1	1	73.8	74.5
37	75.71	461	89	1.16	1.2	83.1	82.2
38	51.91	461	287	1.09	1.24	66.8	70.5
39	51.45	461	291	0.94	-0.71	73.8	70.8
40	42.78	461	359	0.9	-0.73	81.8	79.3
41	39.25	461	381	1.08	0.54	84.4	83.3
42	50.18	461	302	0.97	-0.29	72.9	71.8
43	33.31	461	410	0.66	-1.76	89.2	89.1
44	58.37	461	228	1.04	0.62	69.4	68.9
45	61.53	461	199	1.22	3.31	66.2	69.7
46	50.53	461	299	1.15	1.79	70.5	71.5
47	46.37	461	333	0.88	-1.22	75.9	75.3
48	54.03	461	268	1.04	0.68	69.8	69.5
49	40.44	461	374	0.77	-1.66	81.6	82
50	30.91	461	419	0.65	-1.59	91.3	91

Table 8.
Grade 7 Fall 2006 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	15.79	400	386	0.73	-0.69	96.5	96.5
2	45.01	400	262	0.8	-2.74	70.8	70.2
3	55.85	400	173	1.07	1.22	67.5	68.1
4	34.81	400	329	0.89	-0.7	83	82.5
5	49.49	400	226	0.8	-3.47	75.8	67.4
6	54.52	400	184	0.98	-0.43	67.5	67.7
7	47.42	400	243	0.72	-4.56	77.3	68.3
8	60.13	400	139	0.98	-0.21	74.3	70.9
9	47.29	400	244	1.02	0.31	66	68.4
10	46.04	400	254	0.95	-0.63	71.3	69.3
11	40.47	400	295	0.78	-2.24	78.5	75.2
12	48.28	400	236	0.84	-2.67	72.3	67.9
13	40.17	400	297	0.78	-2.25	76.5	75.6
14	58.33	400	153	1.13	1.97	68.3	69.5
15	49.85	400	223	0.95	-0.85	67	67.4
16	43.7	400	272	0.87	-1.59	73.3	71.4
17	35.74	400	324	0.84	-1.19	81.8	81.3
18	54.52	400	184	1.37	5.96	57	67.7
19	39.25	400	303	0.83	-1.56	78	76.8
20	52.13	400	204	0.83	-3.21	71.8	67.2
21	65.36	400	102	1.29	2.76	76.8	76.8
22	57.95	400	156	1.1	1.58	63.5	69.2
23	62.43	400	122	1.05	0.69	75.3	73.2
24	49.97	400	222	1.08	1.29	66.8	67.3
25	37.15	400	316	1.03	0.28	80.3	79.5
26	30.26	400	350	0.79	-1.18	88	87.5
27	51.53	400	209	1.09	1.55	63	67.2
28	60.26	400	138	1.34	4.34	64	71
29	61.87	400	126	1.16	2	67.8	72.6
30	53.8	400	190	1.27	4.45	55.8	67.5
31	52.85	400	198	1.09	1.59	60.5	67.3
32	51.05	400	213	0.85	-2.77	75	67.2
33	63.42	400	115	1.01	0.16	76.5	74.4
34	40.17	400	297	0.78	-2.24	77.5	75.6
35	54.16	400	187	1.08	1.42	64.5	67.6
36	33.63	400	335	0.67	-2.35	85	83.9
37	53.09	400	196	0.95	-0.81	68	67.3
38	46.29	400	252	0.94	-0.82	70.8	69.1
39	60.39	400	137	1.03	0.38	73.3	71.2
40	60.93	400	133	1.28	3.54	64	71.7
41	47.66	400	241	0.99	-0.1	64	68.2
42	64.75	400	106	1.17	1.73	77.8	76
43	61.74	400	127	1.35	4.14	67	72.5
44	63.28	400	116	1.16	1.82	73.3	74.2
45	45.14	400	261	0.83	-2.26	76.5	70.1
46	44.49	400	266	0.95	-0.61	73.3	70.6
47	56.1	400	171	0.98	-0.25	69	68.2
48	54.52	400	184	1.03	0.54	70	67.7
49	64.15	400	110	1.19	1.99	75.3	75.2
50	32.79	400	339	0.84	-0.98	85	84.9

Table 9.
Grade 8 Fall 2006 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	35.3	441	366	1.37	2.29	83.2	83.6
2	53.11	441	232	1.25	4.07	63.3	68.7
3	56.25	441	204	0.97	-0.46	67.3	69
4	50.5	441	255	0.88	-1.98	71.4	69.3
5	58.41	441	185	0.88	-1.89	77.3	69.8
6	44.75	441	303	0.73	-3.59	74.4	73.3
7	45.26	441	299	0.72	-3.78	75.3	72.8
8	50.39	441	256	1.13	1.99	63	69.4
9	51.19	441	249	1	-0.05	67.3	69.1
10	28.34	441	397	0.97	-0.07	89.6	90.2
11	45.51	441	297	0.86	-1.77	75.7	72.6
12	36.39	441	360	0.68	-2.53	83.4	82.4
13	44.11	441	308	0.96	-0.44	72.6	73.9
14	55.91	441	207	1.1	1.74	65.3	68.9
15	51.76	441	244	0.9	-1.76	72.1	68.9
16	39.79	441	339	0.81	-1.71	79.1	78.6
17	33.96	441	373	0.77	-1.54	85.3	85.1
18	51.07	441	250	1.39	5.7	58.5	69.1
19	28.89	441	395	1.2	0.96	89.8	89.7
20	47.59	441	280	0.85	-2.13	73.5	70.9
21	60.15	441	170	0.95	-0.71	75.3	70.6
22	61.94	441	155	1.07	0.92	72.8	71.8
23	51.53	441	246	1.03	0.49	69.8	69
24	43.58	441	312	0.75	-3.03	77.6	74.5
25	36.21	441	361	0.85	-1.09	85	82.6
26	43.05	441	316	1.03	0.32	75.5	75
27	48.06	441	276	0.93	-1.01	70.7	70.6
28	51.19	441	249	0.83	-3.02	74.1	69.1
29	29.43	441	393	0.7	-1.57	89.8	89.3
30	52.09	441	241	1.01	0.13	68.3	68.8
31	66.86	441	117	1.31	2.84	76.2	76.2
32	53.33	441	230	1.25	3.96	59.2	68.7
33	48.3	441	274	0.86	-2.17	75.1	70.5
34	52.55	441	237	1.11	1.91	66.4	68.7
35	60.15	441	170	0.86	-2.12	73.9	70.6
36	70.43	441	93	1.91	5.72	75.7	80
37	44.11	441	308	1.04	0.53	75.3	73.9
38	77.41	441	56	2.39	5.45	87.5	87.4
39	61.58	441	158	1.22	2.74	68.5	71.6
40	56.03	441	206	1.18	3	61	68.9
41	49.35	441	265	0.94	-0.98	73.2	69.8
42	61.7	441	157	1.16	2.08	68.7	71.7
43	47.11	441	284	0.94	-0.78	70.1	71.3
44	53.56	441	228	0.93	-1.15	73.7	68.7
45	60.74	441	165	1.2	2.7	68.7	71
46	41.68	441	326	0.67	-3.65	81.6	76.5
47	47.47	441	281	0.91	-1.28	72.8	71
48	56.03	441	206	1	0.08	66.4	68.9
49	60.62	441	166	0.93	-0.94	75.3	71
50	45.26	441	299	1.07	0.88	72.1	72.8

Table 10.
Grade 1 Winter 2007 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	15.18	1301	1261	1.02	0.18	97.3	97
2	29.8	1301	1178	1.6	2.79	91.2	91.4
3	31.57	1301	1162	1.51	2.61	91.8	90.5
4	22.16	1301	1231	0.92	-0.24	95	94.9
5	34.31	1301	1134	1.02	0.17	90.2	88.9
6	47.77	1301	935	0.76	-3.35	83.6	79.8
7	51.14	1301	869	0.92	-1.17	81.9	77.5
8	46.57	1301	957	1.38	4.12	79	80.6
9	53.48	1301	820	0.88	-2.06	80.5	76
10	35.73	1301	1118	0.95	-0.3	89.4	88
11	48.72	1301	917	0.8	-2.83	84.8	79.1
12	52.49	1301	841	0.8	-3.33	80.7	76.6
13	57.23	1301	737	1.08	1.44	71	73.9
14	55.81	1301	769	0.92	-1.41	77.2	74.7
15	56.39	1301	756	1.19	3.16	69.3	74.3
16	58.19	1301	715	0.77	-4.47	80.6	73.5
17	52.35	1301	844	0.77	-3.79	82.1	76.7
18	57.05	1301	741	1	0.1	75.3	74
19	59.36	1301	688	0.76	-4.52	78.6	73.1
20	73.31	1301	373	0.77	-2.7	78	76.9
21	69.63	1301	451	0.91	-1.17	80.8	74.4
22	74.92	1301	341	0.85	-1.64	83.4	78.4
23	73.06	1301	378	0.75	-3.09	80.9	76.8
24	84.21	1301	189	0.77	-1.68	88.7	86.5
25	36.4	1301	1110	1.23	1.54	88.5	87.5
26	40.26	1301	1059	1.77	5.47	82.3	84.9
27	62.01	1301	626	1.73	9.9	64.6	72.5
28	63.42	1301	593	1.86	9.9	62.3	72.4
29	36.32	1301	1111	2.15	6.23	84.2	87.6
30	21.2	1301	1236	3.39	6.1	94.6	95.3

Table 11.
Grade 2 Winter 2007 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	30.57	1387	1287	0.67	-2.57	92.9	92.8
2	47.62	1387	1030	0.88	-2.14	78.6	77
3	41.37	1387	1153	0.77	-2.95	84.3	84
4	38.48	1387	1198	0.73	-2.97	86.9	86.8
5	51.81	1387	929	0.87	-3.03	74.9	72.2
6	53.37	1387	888	0.86	-3.73	76	70.7
7	38.06	1387	1204	0.79	-2.26	88.2	87.2
8	57.28	1387	780	1	-0.13	69.3	67.8
9	28.25	1387	1305	0.82	-1.17	94.1	94.1
10	44.59	1387	1094	0.9	-1.5	81	80.5
11	56.36	1387	806	0.83	-5.1	73.9	68.3
12	44.84	1387	1089	0.8	-3.11	80.5	80.2
13	60.7	1387	682	0.94	-1.99	68.1	67
14	65.34	1387	550	1.07	1.96	72.5	68.7
15	40.27	1387	1171	0.73	-3.33	86.5	85.1
16	66.61	1387	515	1.22	5.11	73.4	69.8
17	49.72	1387	981	0.77	-4.96	78.4	74.5
18	70.26	1387	420	1.15	2.85	75.3	73.4
19	60.35	1387	692	0.9	-3.05	69.9	67
20	32.34	1387	1271	0.82	-1.43	92.1	91.7
21	34.64	1387	1247	0.7	-2.8	90.4	90.1
22	53.63	1387	881	1.11	2.6	66.7	70.4
23	29.96	1387	1292	1.03	0.28	93.3	93.2
24	40.27	1387	1171	1.07	0.86	85.8	85.1
25	29.33	1387	1297	0.83	-1.12	93.7	93.5
26	32.13	1387	1273	0.8	-1.56	92.1	91.8
27	75.47	1387	302	1.78	9.49	75.1	79.6
28	58.65	1387	741	1.31	8.64	56.1	67.3
29	60.84	1387	678	1.23	6.69	59.3	67
30	61.77	1387	651	1.08	2.4	64.7	67.1
31	64.44	1387	575	1.01	0.31	68.6	68.2
32	72.69	1387	362	1.25	4.08	76.4	76.2
33	67.92	1387	480	1.01	0.19	71.1	71
34	36.64	1387	1223	0.91	-0.79	88.3	88.5
35	39.5	1387	1183	1.24	2.42	85.8	85.9
36	48.75	1387	1004	1	0.09	75	75.6
37	53.07	1387	896	0.86	-3.5	73.5	70.9
38	60.38	1387	691	1.29	8.37	58.1	67
39	55.68	1387	825	1.2	5.13	62.9	68.7
40	60.73	1387	681	1.24	6.99	63.4	67
41	47.84	1387	1025	0.86	-2.59	78.5	76.7
42	50.25	1387	968	1.01	0.23	75	73.9
43	69.7	1387	434	1.21	4.08	72.3	72.8
44	39.95	1387	1176	1.09	1.03	85	85.4
45	27.53	1387	1310	0.95	-0.26	94.5	94.5

Table 12.
Grade 3 Winter 2007 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	30.84	1358	1242	0.83	-1.1	91.6	91.6
2	48.59	1358	956	0.85	-2.76	75.4	75.2
3	59.84	1358	669	0.92	-2.07	75	69.5
4	57.71	1358	726	0.98	-0.47	69.1	69.6
5	26.68	1358	1276	0.72	-1.62	93.9	94.1
6	51.49	1358	887	0.8	-4.38	75.3	72.7
7	43.46	1358	1064	0.86	-1.89	81.7	80.2
8	56.09	1358	769	0.76	-6.47	78.2	69.9
9	63	1358	585	1.03	0.77	72.8	70.2
10	30.08	1358	1249	0.84	-0.99	92.3	92.1
11	41.61	1358	1098	0.78	-2.81	83.5	82.1
12	34.15	1358	1207	0.75	-2.07	89.5	89.1
13	36.97	1358	1171	0.53	-5.11	88.1	86.7
14	70.04	1358	410	1.04	0.79	76.7	75.2
15	62.36	1358	602	0.95	-1.29	74.2	70
16	41.44	1358	1101	0.99	-0.06	82.8	82.3
17	45.08	1358	1032	0.91	-1.29	78.2	78.6
18	35.39	1358	1192	0.82	-1.56	88.7	88.1
19	53.28	1358	842	1.15	3.21	65.5	71.4
20	53.48	1358	837	0.97	-0.71	71.8	71.3
21	52.41	1358	864	1.11	2.34	67.4	72
22	59.95	1358	666	0.94	-1.54	70.5	69.5
23	39.46	1358	1134	0.77	-2.54	85.7	84.3
24	56.36	1358	762	0.95	-1.27	70.5	69.8
25	51.24	1358	893	1.14	2.76	71.2	72.8
26	63.19	1358	580	1.14	3.27	66.6	70.3
27	69.35	1358	426	0.96	-0.66	76.7	74.6
28	40.56	1358	1116	1.14	1.52	83.6	83.2
29	41.72	1358	1096	0.89	-1.26	83.1	82
30	42.61	1358	1080	0.78	-2.95	82.6	81.1
31	49.02	1358	946	1.31	5.11	72.6	74.7
32	45.23	1358	1029	1.58	7.1	73.7	78.5
33	59.62	1358	675	1.1	2.41	66.7	69.5
34	57.94	1358	720	0.93	-1.84	70.4	69.5
35	37.26	1358	1167	0.9	-0.88	86.2	86.4
36	42	1358	1091	0.99	-0.03	81.8	81.7
37	72.62	1358	353	1.43	5.87	72.5	77.6
38	68.5	1358	446	1.27	4.8	69.2	73.8
39	44.44	1358	1045	1.19	2.49	78.9	79.3
40	50.83	1358	903	0.84	-3.27	75.6	73.2
41	45.86	1358	1016	0.85	-2.37	81.1	77.8
42	55.6	1358	782	1.07	1.6	68.9	70.1
43	38.42	1358	1150	1	0.05	84.7	85.3
44	49.96	1358	924	1.05	0.94	72.6	73.9
45	57.03	1358	744	1.23	5.35	62.2	69.7
46	49.15	1358	943	0.76	-4.59	78.6	74.6
47	51.85	1358	878	0.92	-1.65	72.6	72.4
48	66.78	1358	488	1.61	9.9	61.9	72.5
49	67.31	1358	475	1.12	2.36	69.9	72.8
50	32.16	1358	1229	0.86	-1	91.2	90.7

Table 13.
Grade 4 Winter 2007 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	30.84	1358	1242	0.83	-1.1	91.6	91.6
2	48.59	1358	956	0.85	-2.76	75.4	75.2
3	59.84	1358	669	0.92	-2.07	75	69.5
4	57.71	1358	726	0.98	-0.47	69.1	69.6
5	26.68	1358	1276	0.72	-1.62	93.9	94.1
6	51.49	1358	887	0.8	-4.38	75.3	72.7
7	43.46	1358	1064	0.86	-1.89	81.7	80.2
8	56.09	1358	769	0.76	-6.47	78.2	69.9
9	63	1358	585	1.03	0.77	72.8	70.2
10	30.08	1358	1249	0.84	-0.99	92.3	92.1
11	41.61	1358	1098	0.78	-2.81	83.5	82.1
12	34.15	1358	1207	0.75	-2.07	89.5	89.1
13	36.97	1358	1171	0.53	-5.11	88.1	86.7
14	70.04	1358	410	1.04	0.79	76.7	75.2
15	62.36	1358	602	0.95	-1.29	74.2	70
16	41.44	1358	1101	0.99	-0.06	82.8	82.3
17	45.08	1358	1032	0.91	-1.29	78.2	78.6
18	35.39	1358	1192	0.82	-1.56	88.7	88.1
19	53.28	1358	842	1.15	3.21	65.5	71.4
20	53.48	1358	837	0.97	-0.71	71.8	71.3
21	52.41	1358	864	1.11	2.34	67.4	72
22	59.95	1358	666	0.94	-1.54	70.5	69.5
23	39.46	1358	1134	0.77	-2.54	85.7	84.3
24	56.36	1358	762	0.95	-1.27	70.5	69.8
25	51.24	1358	893	1.14	2.76	71.2	72.8
26	63.19	1358	580	1.14	3.27	66.6	70.3
27	69.35	1358	426	0.96	-0.66	76.7	74.6
28	40.56	1358	1116	1.14	1.52	83.6	83.2
29	41.72	1358	1096	0.89	-1.26	83.1	82
30	42.61	1358	1080	0.78	-2.95	82.6	81.1
31	49.02	1358	946	1.31	5.11	72.6	74.7
32	45.23	1358	1029	1.58	7.1	73.7	78.5
33	59.62	1358	675	1.1	2.41	66.7	69.5
34	57.94	1358	720	0.93	-1.84	70.4	69.5
35	37.26	1358	1167	0.9	-0.88	86.2	86.4
36	42	1358	1091	0.99	-0.03	81.8	81.7
37	72.62	1358	353	1.43	5.87	72.5	77.6
38	68.5	1358	446	1.27	4.8	69.2	73.8
39	44.44	1358	1045	1.19	2.49	78.9	79.3
40	50.83	1358	903	0.84	-3.27	75.6	73.2
41	45.86	1358	1016	0.85	-2.37	81.1	77.8
42	55.6	1358	782	1.07	1.6	68.9	70.1
43	38.42	1358	1150	1	0.05	84.7	85.3
44	49.96	1358	924	1.05	0.94	72.6	73.9
45	57.03	1358	744	1.23	5.35	62.2	69.7
46	49.15	1358	943	0.76	-4.59	78.6	74.6
47	51.85	1358	878	0.92	-1.65	72.6	72.4
48	66.78	1358	488	1.61	9.9	61.9	72.5
49	67.31	1358	475	1.12	2.36	69.9	72.8
50	32.16	1358	1229	0.86	-1	91.2	90.7

Table 14.
Grade 5 Winter 2007 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	49.21	1431	961	0.8	-4.39	75.1	72.2
2	60.01	1431	655	0.93	-2.17	71.4	68.5
3	40.61	1431	1158	0.84	-2.17	81.9	81.6
4	39.2	1431	1184	1.04	0.44	83.8	83.2
5	38.69	1431	1193	0.94	-0.65	84.7	83.8
6	53.72	1431	837	0.83	-4.67	74.8	69.1
7	48.13	1431	989	1.2	3.66	73.8	73.3
8	61.58	1431	610	1.1	2.84	70.3	69.2
9	58.18	1431	708	0.92	-2.54	71.9	68.2
10	48.36	1431	983	0.89	-2.27	74.8	73
11	53.4	1431	846	0.84	-4.42	74.6	69.3
12	37.32	1431	1216	0.7	-3.49	85.6	85.3
13	48.48	1431	980	0.76	-5.31	77.5	72.9
14	45.22	1431	1060	0.86	-2.49	76.5	76.3
15	37.01	1431	1221	0.91	-0.87	85	85.6
16	48.05	1431	991	0.98	-0.33	74.8	73.3
17	45.52	1431	1053	1.04	0.66	76.5	76
18	53.4	1431	846	1	0.12	71.4	69.3
19	63.8	1431	548	0.99	-0.26	75.5	70.6
20	45.64	1431	1050	0.96	-0.61	77.4	75.8
21	35.56	1431	1243	1.09	0.86	87	87
22	45.98	1431	1042	0.9	-1.8	76.6	75.5
23	36.43	1431	1230	0.75	-2.69	86.7	86.2
24	51.69	1431	894	1.29	6.4	63.8	70.3
25	41.12	1431	1148	1.08	1.01	81.6	81
26	70.88	1431	369	1.65	9.44	71.8	77.4
27	69.58	1431	399	1.29	5.02	73.5	76
28	58.98	1431	685	1.05	1.57	66.5	68.3
29	60.74	1431	634	1.14	3.92	66.2	68.8
30	47.93	1431	994	0.88	-2.41	76.7	73.4
31	60.92	1431	629	1.16	4.36	65.1	68.9
32	36.75	1431	1225	0.76	-2.63	86.4	85.9
33	51	1431	913	1.09	2.15	67.1	70.8
34	60.32	1431	646	1.08	2.37	65	68.7
35	58.11	1431	710	0.98	-0.61	69.3	68.2
36	62.08	1431	596	1.06	1.77	67.1	69.5
37	62.57	1431	582	1.32	8.01	62.9	69.8
38	55.22	1431	794	1.05	1.32	66.5	68.5
39	52.34	1431	876	0.85	-4	72.5	69.9
40	51.18	1431	908	1.07	1.69	67.9	70.7
41	35.42	1431	1245	1.06	0.63	87.6	87.2
42	55.22	1431	794	1	-0.06	68.9	68.5
43	33.9	1431	1266	0.91	-0.78	88.9	88.6
44	40.13	1431	1167	0.88	-1.45	83.9	82.2
45	64.57	1431	527	1.07	1.82	71.1	71.2
46	49.36	1431	957	1.07	1.51	71.4	72.1
47	45.98	1431	1042	1.02	0.44	76	75.5
48	43.67	1431	1095	0.95	-0.76	79.4	78.1
49	48.28	1431	985	1.04	0.84	71.3	73.1
50	38.57	1431	1195	1.08	0.85	84.6	83.9

Table 15.
Grade 6 Winter 2007 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	31.59	483	435	1.06	0.35	90.3	90.1
2	54.77	483	272	1.02	0.41	69.2	69.6
3	58.3	483	238	1.13	2.09	66.3	68.9
4	79.58	483	69	1.95	4.5	86.5	86.3
5	55.09	483	269	0.86	-2.35	72.3	69.5
6	63.31	483	190	0.93	-1.07	74.7	70.1
7	43.9	483	366	1.15	1.35	74.7	78
8	41.48	483	383	0.99	-0.03	80.7	80.5
9	52.66	483	292	0.87	-2.05	75.6	70.7
10	49.45	483	321	0.89	-1.41	75.4	73
11	44.43	483	362	0.7	-3.07	80.1	77.5
12	53.93	483	280	1.04	0.63	67.1	70
13	40.72	483	388	0.64	-3.01	84.3	81.4
14	53.08	483	288	0.82	-2.87	73.5	70.4
15	30.52	483	439	0.79	-0.89	90.7	90.9
16	43.76	483	367	0.95	-0.44	79.1	78.2
17	68.74	483	142	1.19	2.06	69.2	74.4
18	48.53	483	329	1.04	0.45	71.4	73.7
19	50.91	483	308	1.05	0.74	70.8	71.8
20	39.78	483	394	0.83	-1.24	84.1	82.4
21	51.35	483	304	1.4	4.99	60.5	71.5
22	42.22	483	378	0.73	-2.41	81.8	79.8
23	42.93	483	373	0.88	-1.02	76.6	79
24	33.3	483	428	1.08	0.45	89.2	88.7
25	38.61	483	401	0.94	-0.39	85.7	83.6
26	30.52	483	439	0.89	-0.44	90.7	90.9
27	59.64	483	225	1.07	1.24	66.7	69.1
28	54.88	483	271	0.92	-1.32	71.8	69.5
29	48.64	483	328	1.1	1.17	75.8	73.6
30	60.26	483	219	1.29	4.53	61.3	69.1
31	45.35	483	355	0.9	-0.98	78.9	76.6
32	33.06	483	429	0.79	-1.05	89	88.9
33	53.72	483	282	1.22	3.27	63.1	70.1
34	59.54	483	226	0.89	-1.85	73.9	69
35	59.02	483	231	1.11	1.77	67.7	69
36	44.83	483	359	0.76	-2.41	78.5	77.1
37	43.35	483	370	1.09	0.76	76.8	78.6
38	52.66	483	292	0.97	-0.36	69.8	70.7
39	44.03	483	365	0.83	-1.6	80.3	77.9
40	50.24	483	314	1.01	0.11	73.1	72.3
41	59.33	483	228	1.21	3.42	65	69
42	35.86	483	416	0.72	-1.72	86.5	86.4
43	65.72	483	168	1.14	1.83	68.7	71.7
44	63.74	483	186	0.99	-0.14	71.4	70.4
45	33.06	483	429	0.64	-1.97	89.4	88.9
46	47.22	483	340	0.82	-2.06	79.5	74.9
47	61.93	483	203	1.15	2.27	65	69.5
48	77.38	483	81	1.56	3.26	82.4	84.1
49	66.4	483	162	1.25	2.97	70.4	72.2
50	36.64	483	412	0.67	-2.17	86.5	85.6

Table 16.
Grade 7 Winter 2007 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	27.65	425	390	1.1	0.45	92	91.8
2	47.31	425	282	0.86	-1.56	74.6	71.5
3	58.12	425	191	1.16	2.48	63.5	69.2
4	50.27	425	258	1.08	1	67.1	69.7
5	53.45	425	231	0.85	-2.42	73.9	68.6
6	52.28	425	241	1.08	1.15	67.3	68.9
7	59.91	425	176	1.19	2.74	65.4	70.1
8	36.31	425	355	0.9	-0.56	84	83.7
9	58.24	425	190	1.26	3.79	65.2	69.2
10	47.94	425	277	1.24	2.62	72.9	71
11	43.59	425	310	0.71	-2.82	77.2	75.1
12	40.48	425	331	0.81	-1.44	79.5	78.7
13	58.95	425	184	0.93	-1.04	70.4	69.5
14	48.69	425	271	0.84	-2.02	72	70.5
15	51.22	425	250	0.92	-1.15	69.6	69.2
16	41.41	425	325	0.88	-0.94	80.5	77.6
17	36.31	425	355	0.8	-1.2	84	83.7
18	66.74	425	123	1.32	3.23	71.3	75.7
19	55.66	425	212	1.12	1.79	65.6	68.6
20	46.8	425	286	0.95	-0.47	69.6	71.9
21	39.34	425	338	1.17	1.19	78.8	80.1
22	42.74	425	316	0.7	-2.76	77.4	76
23	59.07	425	183	0.76	-4.03	78.6	69.6
24	53.22	425	233	0.92	-1.22	72.9	68.6
25	33.61	425	368	0.92	-0.37	87.3	86.7
26	34.27	425	365	0.88	-0.6	86.1	86
27	44.7	425	302	1.02	0.18	74.8	73.9
28	55.43	425	214	0.95	-0.86	69.4	68.6
29	61.87	425	160	1.18	2.47	71.8	71.4
30	48.19	425	275	0.85	-1.84	74.4	70.8
31	66.6	425	124	1.31	3.15	69.6	75.5
32	32.2	425	374	0.88	-0.55	88.5	88.1
33	42.45	425	318	0.86	-1.22	78.4	76.4
34	66.32	425	126	1.49	4.78	69.9	75.3
35	48.93	425	269	0.91	-1.11	71.1	70.4
36	51.58	425	247	0.94	-0.88	70.4	69.1
37	50.27	425	258	0.99	-0.17	70.8	69.7
38	47.69	425	279	0.97	-0.34	71.5	71.2
39	65.9	425	129	0.89	-1.23	77.6	74.9
40	50.14	425	259	1.02	0.31	68.2	69.7
41	69.86	425	102	1.07	0.67	79.3	79
42	48.93	425	269	0.76	-3.11	73.4	70.4
43	43.45	425	311	1	0.01	75.1	75.2
44	43.03	425	314	0.74	-2.42	78.8	75.7
45	54.62	425	221	1.05	0.75	67.1	68.5
46	61.99	425	159	1.32	4.06	69.6	71.5
47	52.52	425	239	0.93	-1.03	75.3	68.8
48	40.32	425	332	0.87	-0.92	80.2	78.9
49	61.87	425	160	1.27	3.58	66.6	71.4
50	47.57	425	280	1.05	0.6	70.8	71.3

Table 17.
Grade 8 Winter 2007 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	25.71	443	402	1.05	0.32	90.5	90.7
2	23.87	443	408	0.64	-1.71	92.1	92.1
3	53.9	443	202	1.04	0.71	66.1	67.5
4	46.92	443	266	0.86	-2.48	72.9	68.7
5	41	443	316	0.95	-0.51	73.6	73.6
6	40.87	443	317	0.82	-2.14	75.2	73.8
7	55.99	443	183	1.01	0.13	69.3	68.4
8	46.25	443	272	1.06	1.02	67.5	69.1
9	50.54	443	233	0.98	-0.35	68.8	67.4
10	43.94	443	292	0.98	-0.2	70.4	70.7
11	33.8	443	365	0.72	-2.27	82.4	82.6
12	60.14	443	147	1.08	1.07	71.3	71.8
13	63.56	443	120	1.16	1.79	77.4	75.3
14	46.7	443	268	0.83	-3	74.7	68.8
15	78.48	443	40	2.32	4.66	90.7	91
16	34.98	443	358	0.76	-2.01	80.1	81.1
17	27.34	443	396	0.71	-1.59	89.6	89.4
18	56.21	443	181	0.84	-3.1	74.7	68.5
19	66.68	443	98	1.82	6.26	77	79.1
20	58.37	443	162	0.86	-2.27	75.4	70.1
21	67.29	443	94	1.18	1.55	80.1	79.9
22	38.32	443	336	1.04	0.38	76.3	76.8
23	58.02	443	165	1.15	2.41	67.9	69.9
24	57.33	443	171	1.12	1.91	68.2	69.3
25	53.79	443	203	1.37	6.5	57.8	67.5
26	60.38	443	145	1.02	0.31	71.8	72
27	69.59	443	80	1.36	2.59	81.5	82.6
28	44.06	443	291	0.86	-2.15	75.4	70.6
29	30.78	443	381	0.72	-1.9	86	86.1
30	23.87	443	408	0.54	-2.28	92.1	92.1
31	44.3	443	289	0.86	-2.09	74.5	70.5
32	33.09	443	369	0.65	-2.82	84.4	83.4
33	55.32	443	189	1.09	1.72	64.1	68.1
34	57.79	443	167	1.21	3.28	64.8	69.7
35	59.3	443	154	1.13	1.93	66.4	71
36	42.38	443	305	0.92	-1	72.5	72.2
37	53.68	443	204	0.96	-0.73	71.6	67.5
38	43.23	443	298	0.87	-1.81	75.4	71.4
39	58.71	443	159	1.13	2.02	67.9	70.4
40	65.79	443	104	1.42	3.72	77	78
41	58.71	443	159	0.8	-3.29	77	70.4
42	58.71	443	159	1.06	0.97	67.9	70.4
43	49.13	443	246	0.9	-2.03	68.6	67.7
44	40.09	443	323	0.85	-1.77	79.7	74.7
45	39.56	443	327	0.83	-1.88	80.1	75.3
46	58.37	443	162	0.93	-1.08	71.8	70.1
47	46.25	443	272	1.1	1.64	67.9	69.1
48	52.49	443	215	0.95	-1.08	69.1	67.3
49	59.42	443	153	1.17	2.37	68.4	71.1
50	60.02	443	148	1.23	3.07	67	71.6
51	63.56	443	120	1.24	2.58	73.4	75.3
52	62.38	443	129	0.92	-0.93	76.5	74
53	39.69	443	326	0.94	-0.65	77.7	75.1
54	54.55	443	196	0.94	-1.14	70.7	67.7
55	34.82	443	359	0.84	-1.34	81.9	81.3

Table 18.
Grade 1 Spring 07 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	10.07	1376	1366	0.27	-3.14	99.3	99.3
2	28.77	1376	1327	0.9	-0.32	97.2	96.7
3	22.36	1376	1347	0.52	-1.85	98.2	98
4	47.5	1376	1179	0.77	-1.98	89	87.6
5	37.97	1376	1276	1.22	1.21	93.8	93.4
6	33.11	1376	1307	1.04	0.27	95.9	95.4
7	53.86	1376	1080	0.94	-0.66	83	82.2
8	62.19	1376	905	0.88	-2.16	76.6	74.3
9	50.57	1376	1135	0.91	-0.86	85.8	85.1
10	34.86	1376	1297	0.79	-1.02	95.2	94.8
11	56.39	1376	1032	0.88	-1.58	80.1	79.8
12	68.87	1376	737	0.81	-3.93	74.7	70.4
13	54.78	1376	1063	0.99	-0.09	82.3	81.3
14	55.57	1376	1048	0.99	-0.16	81.8	80.5
15	92.13	1376	216	0.97	-0.23	88	85.6
16	40.78	1376	1253	1.01	0.12	92.4	92
17	60.29	1376	949	0.98	-0.34	76.2	75.9
18	80.82	1376	433	0.95	-0.76	79.1	75
19	87.09	1376	301	0.89	-1.2	82.6	81
20	27.15	1376	1333	1.09	0.43	97.3	97.1
21	38.5	1376	1272	0.67	-2.05	93.5	93.1
22	43.07	1376	1231	0.59	-3.13	91.6	90.6
23	36.72	1376	1285	0.89	-0.54	94	94
24	62.49	1376	898	0.91	-1.53	74.1	74
25	69.29	1376	726	1.87	9.9	58.4	70.2
26	34.86	1376	1297	2.05	3.96	93.6	94.8
27	23.16	1376	1345	1.16	0.61	97.5	97.9
28	44.47	1376	1216	1.12	0.87	87.8	89.7
29	53.58	1376	1085	1.22	2.36	80	82.4
30	88.74	1376	271	5.47	9.91	78.1	82.6

Table 19.
Grade 2 Spring 07 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	45.71	1448	1177	1.03	0.39	82.4	82.6
2	59.12	1448	862	0.97	-0.61	72.5	71
3	52.34	1448	1037	0.79	-3.8	79.6	75.9
4	32.77	1448	1347	1.1	0.63	93.2	93.1
5	54.54	1448	983	0.78	-4.6	77.5	74
6	45.44	1448	1182	0.81	-2.3	82.6	82.9
7	33.35	1448	1342	0.79	-1.34	93	92.8
8	68.93	1448	587	0.93	-1.62	77.8	71.2
9	60.95	1448	811	0.9	-2.62	72.7	70.3
10	56.25	1448	939	0.8	-4.59	76.7	72.7
11	49.93	1448	1092	0.7	-5.04	82.1	78.2
12	42.54	1448	1231	1.01	0.13	85.2	85.7
13	35.36	1448	1323	1.08	0.59	91.9	91.5
14	58.09	1448	890	0.72	-7.29	78.9	71.5
15	54.62	1448	981	0.81	-3.98	77.6	73.9
16	58.35	1448	883	0.83	-4.23	75.3	71.4
17	62.2	1448	776	0.9	-2.86	76.1	70
18	42.86	1448	1226	0.75	-2.67	86.4	85.4
19	46.67	1448	1159	0.73	-3.6	83.6	81.6
20	24.68	1448	1398	0.56	-2.19	96.6	96.6
21	36.68	1448	1309	1.07	0.5	91	90.6
22	36.32	1448	1313	0.98	-0.08	91.4	90.9
23	67.84	1448	617	1.14	3.46	66.4	70.8
24	47.34	1448	1146	1	-0.01	81.4	80.9
25	53.9	1448	999	0.98	-0.3	75.5	74.5
26	42.67	1448	1229	1.6	5.07	85	85.6
27	33	1448	1345	0.73	-1.74	93	93
28	68.17	1448	608	1.07	1.65	69.3	70.9
29	44.25	1448	1203	1.15	1.55	83.4	84.1
30	45.66	1448	1178	0.86	-1.72	84.5	82.6
31	52.12	1448	1042	1.02	0.32	76	76.1
32	45.44	1448	1182	1.21	2.27	81.5	82.9
33	51.39	1448	1059	1.42	6.11	71.4	76.8
34	50.74	1448	1074	1.15	2.22	76	77.4
35	62.66	1448	763	1.35	8.42	60.6	69.9
36	64.43	1448	713	1.35	8.49	60.2	69.9
37	57.87	1448	896	1.05	1.13	70.8	71.7
38	69.23	1448	579	1.48	9.9	62	71.4
39	50.96	1448	1069	1.36	5.09	74.3	77.2
40	36.68	1448	1309	0.88	-0.87	90.6	90.6
41	48.9	1448	1114	0.97	-0.39	77.8	79.3
42	59.84	1448	842	0.98	-0.55	70.7	70.7
43	45.71	1448	1177	1.28	3.02	81.3	82.6
44	50.91	1448	1070	0.98	-0.37	76.3	77.2
45	42.61	1448	1230	1.03	0.31	85.6	85.6

Table 20.
Grade 3 Spring 2007 Data File.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	24.67	1396	1331	1.22	1.12	95.1	95.4
2	44.37	1396	1091	0.91	-1.42	80.5	79.9
3	38.82	1396	1189	0.83	-1.94	85.2	85.7
4	49.14	1396	986	0.85	-3.23	77.9	74.9
5	39.02	1396	1186	0.6	-5.24	86.7	85.5
6	53.13	1396	885	0.78	-6.1	76.9	71.3
7	37.32	1396	1211	0.78	-2.36	87.8	87.1
8	46.2	1396	1053	0.69	-6.02	81.1	78
9	49.84	1396	969	0.85	-3.4	78	74.2
10	40.57	1396	1161	1.01	0.15	83.6	84
11	36.96	1396	1216	0.68	-3.56	88.4	87.4
12	56.43	1396	795	0.8	-6.34	75.4	69.1
13	73.93	1396	330	1.01	0.18	78.9	77.7
14	42.83	1396	1121	0.93	-0.97	82.4	81.6
15	47.55	1396	1023	0.75	-5.07	79.4	76.5
16	18.46	1396	1359	0.7	-1.23	97.3	97.3
17	56.25	1396	800	0.97	-1.03	70.2	69.2
18	57.39	1396	768	1.16	4.55	62.7	68.7
19	45.97	1396	1058	0.92	-1.37	78.2	78.2
20	48.5	1396	1001	0.87	-2.71	76.8	75.5
21	42.07	1396	1135	0.86	-1.91	83.5	82.4
22	80.63	1396	205	2.62	9.9	83.7	85.5
23	48.67	1396	997	1.02	0.38	74.6	75.4
24	61.05	1396	664	1.19	5.42	63.8	68
25	47.64	1396	1021	0.89	-2.14	76.7	76.5
26	69.7	1396	429	1.3	5.53	72.9	73.1
27	56.47	1396	794	1.07	2.2	66.6	69.1
28	59.3	1396	714	1.17	5.1	62.3	68.1
29	44.22	1396	1094	0.84	-2.54	80.7	80.1
30	56.21	1396	801	1.11	3.29	66.5	69.2
31	67.94	1396	474	1.1	2.15	72.6	71.5
32	52.6	1396	899	0.83	-4.44	75.3	71.7
33	39.78	1396	1174	0.93	-0.83	84.8	84.8
34	36.66	1396	1220	0.69	-3.34	88	87.7
35	66.99	1396	499	1.3	6.49	67.8	70.7
36	41.62	1396	1143	0.85	-2.1	83.8	82.9
37	39.91	1396	1172	0.83	-2.12	85.4	84.6
38	45.92	1396	1059	0.91	-1.54	78.4	78.3
39	55.57	1396	819	1.14	3.94	64.8	69.6
40	82.1	1396	183	3.36	9.9	86.6	86.9
41	33.83	1396	1255	0.86	-1.17	90.2	90.1
42	28.15	1396	1308	0.7	-2	93.9	93.7
43	50.49	1396	953	1.01	0.16	73.7	73.6
44	46.75	1396	1041	0.89	-1.99	78	77.4
45	34.94	1396	1242	1.18	1.52	89	89.2
46	47.46	1396	1025	1.09	1.68	75.6	76.6
47	75.85	1396	290	1.46	5.61	79.7	80
48	66.24	1396	519	1.27	6.21	63.3	70.2
49	61.54	1396	650	1.1	2.95	67.8	68.1
50	52.33	1396	906	0.91	-2.3	72.5	71.9

Table 21.
Grade 4 Spring 2007 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	30.71	1398	1281	0.93	-0.42	91.8	91.7
2	51.56	1398	907	0.85	-3.38	74	72
3	51.6	1398	906	1.16	3.22	70.2	72
4	41.53	1398	1131	0.76	-3.14	81.8	81.8
5	44.39	1398	1075	1.08	1.13	78.3	78.7
6	43.7	1398	1089	0.8	-2.85	81	79.4
7	42.12	1398	1120	0.93	-0.82	80.8	81.2
8	64.83	1398	547	1.19	4.33	70	71.2
9	40.97	1398	1141	0.83	-2.04	82.6	82.4
10	53.17	1398	865	1	0.09	71.6	71.1
11	66.47	1398	504	1.06	1.24	73.5	72.1
12	48.75	1398	977	0.87	-2.51	75	74.2
13	53.02	1398	869	0.87	-2.98	76	71.1
14	49.61	1398	956	0.84	-3.11	75.7	73.5
15	40.18	1398	1155	0.84	-1.89	84.2	83.3
16	52.29	1398	888	0.87	-3.1	76.5	71.6
17	37.96	1398	1191	0.79	-2.15	85.7	85.6
18	47.95	1398	996	0.86	-2.6	75.1	75
19	49.57	1398	957	0.74	-5.38	76.3	73.5
20	42.9	1398	1105	0.83	-2.28	82.3	80.3
21	23.16	1398	1337	0.64	-1.99	95.8	95.7
22	61.17	1398	646	1.07	1.87	68.2	69.9
23	62.6	1398	607	1.22	5.37	64.8	70.3
24	56.17	1398	784	1.13	3.36	64.4	69.9
25	58.39	1398	723	1.08	2.26	67.9	69.6
26	43.35	1398	1096	0.77	-3.23	82.2	79.8
27	65.85	1398	520	1.28	6.01	68.5	71.7
28	35.59	1398	1225	0.82	-1.58	88.4	87.9
29	39.88	1398	1160	1.14	1.47	84.4	83.6
30	35.06	1398	1232	0.95	-0.42	88.9	88.4
31	54.41	1398	832	1.25	5.67	62.7	70.5
32	49.86	1398	950	1.27	4.82	69	73.3
33	39.65	1398	1164	0.84	-1.77	84.8	83.9
34	52.49	1398	883	1.26	5.27	68.3	71.4
35	49.41	1398	961	0.83	-3.41	75.3	73.6
36	56.79	1398	767	1.08	2.04	66.8	69.8
37	36.39	1398	1214	0.95	-0.4	88	87.2
38	50.18	1398	942	0.82	-3.83	76.6	73
39	57.41	1398	750	0.85	-4.37	74.8	69.7
40	44.82	1398	1066	0.95	-0.71	80.3	78.2
41	58.78	1398	712	1.02	0.62	68.4	69.7
42	45.29	1398	1056	0.89	-1.74	78.7	77.7
43	59.07	1398	704	0.98	-0.5	70.8	69.6
44	50.06	1398	945	0.86	-2.82	74	73.1
45	60.23	1398	672	1.29	7.3	62.4	69.8
46	68.21	1398	460	1.27	5.13	70.8	73.3
47	49.25	1398	965	0.95	-1.01	74.9	73.8
48	65.32	1398	534	1.38	8.11	64.2	71.4
49	65.59	1398	527	1.12	2.72	70.6	71.6
50	52.29	1398	888	1.01	0.3	71.7	71.6

Table 22.
Grade 5 Spring 2007 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	40.72	1490	1219	0.73	-3.51	84.2	82.6
2	42.3	1490	1188	0.67	-4.75	82.9	80.9
3	47.4	1490	1073	0.86	-2.59	78.9	75.6
4	33.92	1490	1327	1.17	1.35	89.1	89.1
5	33.29	1490	1335	0.92	-0.66	89.7	89.7
6	46.2	1490	1102	0.74	-4.65	80.1	76.8
7	47.52	1490	1070	0.96	-0.65	75.4	75.5
8	51.53	1490	965	0.79	-5.04	74.7	72.1
9	46.15	1490	1103	0.68	-5.83	80.6	76.9
10	44.23	1490	1147	0.95	-0.76	79.1	78.9
11	36.75	1490	1287	1.03	0.29	86.2	86.6
12	50.5	1490	993	1.12	2.4	68.8	72.9
13	78.39	1490	259	2.39	9.9	81.5	83.7
14	52.22	1490	946	0.95	-1.04	70.7	71.7
15	49.25	1490	1026	0.93	-1.27	74	73.9
16	43.91	1490	1154	0.82	-2.7	82.3	79.2
17	28.14	1490	1389	0.84	-1.05	93.4	93.2
18	46.99	1490	1083	0.88	-2.15	78.3	76
19	55.31	1490	858	0.77	-6.71	79.6	70.1
20	40.72	1490	1219	0.88	-1.46	82.8	82.6
21	28.03	1490	1390	0.93	-0.39	93.4	93.3
22	58.28	1490	771	1.01	0.4	69.5	69.5
23	67.39	1490	509	1.73	9.9	63.2	72.8
24	45.64	1490	1115	0.92	-1.18	76.4	77.4
25	56.54	1490	822	0.94	-1.74	70.5	69.7
26	56.75	1490	816	1.03	0.75	68.3	69.7
27	47.72	1490	1065	1.22	3.69	71.5	75.3
28	45.38	1490	1121	0.89	-1.77	79.7	77.7
29	41.19	1490	1210	0.81	-2.49	83.7	82.1
30	37.27	1490	1279	1.38	3.34	85.7	86.1
31	65.68	1490	556	1.41	8.87	64.3	71.6
32	63.25	1490	625	0.91	-2.49	74.5	70.3
33	44.95	1490	1131	1.08	1.2	76.9	78.1
34	64.08	1490	601	1.22	5.41	67	70.7
35	52.83	1490	929	1.04	1.03	67.7	71.3
36	68.45	1490	481	1.07	1.45	74.3	73.7
37	48	1490	1058	0.91	-1.7	77.8	75.1
38	56.82	1490	814	0.85	-4.47	74.2	69.7
39	63.6	1490	615	1.19	4.76	68.2	70.5
40	56.88	1490	812	0.94	-1.56	71.7	69.6
41	82.32	1490	195	2.12	8.2	86.4	87.5
42	63.04	1490	631	1.13	3.53	67.7	70.2
43	55.79	1490	844	1.03	0.74	70.3	69.9
44	56.85	1490	813	1.04	1.05	67.2	69.7
45	43.17	1490	1170	0.7	-4.56	81.5	80
46	57.84	1490	784	0.87	-4.03	74.3	69.5
47	30.37	1490	1368	0.86	-0.97	92	91.8
48	49.33	1490	1024	0.92	-1.65	74	73.9
49	45.17	1490	1126	0.98	-0.21	77.2	77.9
50	31.95	1490	1351	0.89	-0.78	90.7	90.7

Table 23.
Grade 6 Spring 2007 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	56.48	483	251	0.82	-3.23	74.5	69
2	54.83	483	267	0.87	-2.29	74.9	69.4
3	57.1	483	245	1.04	0.71	67.5	68.9
4	60	483	217	1.06	1.02	67.3	69.2
5	43.13	483	369	0.98	-0.15	79.5	77.9
6	45.47	483	351	0.93	-0.68	77.8	75.4
7	50.36	483	309	1.04	0.58	67.3	71.5
8	54.41	483	271	0.82	-3.11	75.4	69.6
9	49.47	483	317	0.9	-1.21	76.6	72
10	46.93	483	339	0.92	-0.92	75.4	74.1
11	39.16	483	396	0.58	-3.33	82.4	82.3
12	46.33	483	344	0.78	-2.47	76.8	74.6
13	44.32	483	360	0.89	-1.07	79.1	76.6
14	55.04	483	265	0.84	-2.74	76	69.4
15	46.21	483	345	1.32	3.07	68.7	74.7
16	39.63	483	393	0.94	-0.38	82	81.8
17	36.39	483	412	0.82	-1.09	85.5	85.4
18	56.18	483	254	1.12	2.05	65.6	69.1
19	55.55	483	260	1.16	2.56	67.9	69.2
20	49.58	483	316	0.82	-2.39	77.6	72
21	40.1	483	390	1.18	1.28	81.8	81.2
22	51.01	483	303	1.02	0.28	66.5	71.1
23	51.98	483	294	0.84	-2.5	75.4	70.6
24	44.96	483	355	0.88	-1.16	77.4	76
25	31.93	483	433	1.58	2.41	89.6	89.6
26	35.03	483	419	0.72	-1.64	86.7	86.8
27	34.21	483	423	0.63	-2.18	88	87.6
28	50.03	483	312	0.8	-2.74	77.6	71.7
29	44.45	483	359	0.85	-1.4	80.1	76.5
30	62.64	483	192	1.38	5.37	63.6	70.3
31	51.77	483	296	1.06	0.84	69.2	70.7
32	61.47	483	203	1.2	3.16	63.6	69.7
33	60.52	483	212	1.26	4.11	59.2	69.4
34	26.3	483	452	0.49	-2.23	93.6	93.6
35	44.19	483	361	0.73	-2.68	78.1	76.8
36	55.55	483	260	0.91	-1.55	70.8	69.2
37	64.61	483	174	1.38	4.96	67.9	71.6
38	60.21	483	215	0.93	-1.28	72.7	69.3
39	50.8	483	305	1.03	0.47	68.9	71.2
40	56.9	483	247	1.01	0.14	67.5	68.9
41	53.68	483	278	1.11	1.74	65	69.8
42	48.68	483	324	0.71	-3.82	78.7	72.6
43	46.33	483	344	1.24	2.36	73.5	74.6
44	48.56	483	325	1.11	1.26	68.5	72.7
45	64.27	483	177	1.22	3.06	70.4	71.4
46	53.58	483	279	0.95	-0.74	72.3	69.9
47	67.72	483	147	1.42	4.53	72.9	74.4
48	53.47	483	280	1.4	5.61	60.9	69.9
49	51.55	483	298	0.84	-2.36	75.8	70.8
50	46.93	483	339	1.01	0.12	72	74.1

Table 24.
Grade 7 Spring 2007 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	12.68	418	407	1.12	0.41	97.4	97.4
2	45.95	418	268	0.81	-2.6	73.9	70.3
3	52.61	418	212	1.06	1.04	66.3	68.2
4	46.57	418	263	0.85	-2.17	75.4	69.8
5	32.67	418	355	0.9	-0.57	84.4	85.1
6	43.92	418	284	0.83	-2.04	75.8	71.9
7	55.07	418	191	1.03	0.6	68.7	68.6
8	40.93	418	306	0.92	-0.74	77	75.1
9	48.5	418	247	0.96	-0.54	74.2	68.8
10	44.05	418	283	0.9	-1.15	73.2	71.7
11	46.32	418	265	0.82	-2.55	75.4	70
12	58.18	418	165	1.17	2.62	64.6	69.9
13	53.43	418	205	0.89	-1.92	72	68.2
14	56.01	418	183	0.9	-1.73	75.1	68.8
15	47.9	418	252	0.83	-2.67	73	69.1
16	38.08	418	325	0.8	-1.7	78	78.6
17	30.92	418	363	0.71	-1.68	87.3	86.9
18	73.47	418	64	2.19	5.76	82.3	85.3
19	50.86	418	227	0.95	-0.91	71.5	68.2
20	43.79	418	285	0.86	-1.66	77.5	72
21	62.37	418	132	0.97	-0.33	77.8	73.4
22	73.04	418	66	1.7	3.83	84.7	84.9
23	51.45	418	222	0.86	-2.59	74.4	68.2
24	48.38	418	248	0.86	-2.16	73.4	68.9
25	29.72	418	368	0.74	-1.36	88.5	88.1
26	31.6	418	360	0.89	-0.6	87.1	86.2
27	63.19	418	126	0.93	-0.82	77.3	74.2
28	38.08	418	325	0.79	-1.79	79.4	78.6
29	46.57	418	263	0.82	-2.57	73.4	69.8
30	32.46	418	356	0.78	-1.37	86.1	85.3
31	53.78	418	202	0.95	-0.83	69.4	68.3
32	59.29	418	156	1.33	4.41	62.9	70.7
33	57.57	418	170	0.96	-0.56	72.5	69.5
34	55.78	418	185	1.12	1.96	64.6	68.8
35	44.17	418	282	1	0.04	75.4	71.6
36	64.87	418	114	1.65	5.84	71.1	76
37	64.59	418	116	1.54	5.09	70.6	75.7
38	58.91	418	159	1.23	3.21	67.9	70.4
39	48.14	418	250	0.87	-2.03	74.4	69
40	52.38	418	214	0.82	-3.34	75.4	68.1
41	67.13	418	99	1.09	0.86	78.9	78.5
42	40.79	418	307	0.8	-2.01	77.8	75.2
43	44.3	418	281	0.89	-1.32	76.1	71.5
44	60.55	418	146	1.5	6	61.7	71.7
45	52.73	418	211	1.05	0.82	66	68.2
46	59.54	418	154	0.94	-0.86	73.4	70.9
47	53.2	418	207	1.08	1.38	65.1	68.2
48	68.43	418	91	1.74	5.25	77.8	79.9
49	37.76	418	327	0.79	-1.77	79.9	79
50	57.33	418	172	1.24	3.62	65.6	69.4

Table 25.
Grade 8 Spring 2007 Data.

Item	Measure	Count	Score	Out. Msq.	Out. ZSTD	Obs. Match	Exp. Match
1	19.72	437	420	0.88	-0.23	96.1	96.1
2	48.86	437	281	1.03	0.41	70.5	70.6
3	54.93	437	228	1.07	1.21	65	68.3
4	42.64	437	329	0.95	-0.49	74.8	76.9
5	32.09	437	387	1.42	1.97	88.3	88.6
6	47.77	437	290	0.91	-1.27	73.5	71.4
7	55.71	437	221	1.01	0.13	69.8	68.2
8	51.66	437	257	1.04	0.69	69.3	69
9	51.32	437	260	0.84	-2.73	74.1	69.2
10	45.37	437	309	0.82	-2.13	74.8	73.8
11	46.15	437	303	0.81	-2.39	73.9	72.9
12	60.91	437	175	1.1	1.6	68.2	69.5
13	45.76	437	306	0.85	-1.83	73.2	73.4
14	64.37	437	146	1.3	3.75	71.4	72.1
15	39.92	437	347	1.35	2.61	78.7	80.1
16	34.18	437	378	0.83	-1.02	87	86.6
17	29.96	437	395	0.55	-2.43	90.2	90.4
18	52.35	437	251	0.9	-1.66	71.2	68.8
19	48.5	437	284	0.95	-0.74	72.1	70.8
20	58.85	437	193	0.9	-1.89	76.4	68.7
21	63.51	437	153	0.94	-0.83	74.4	71.3
22	47.28	437	294	0.86	-1.83	74.8	71.8
23	58.17	437	199	0.79	-4.16	79.4	68.5
24	41.62	437	336	0.74	-2.53	81	78.1
25	31.84	437	388	0.91	-0.43	88.6	88.8
26	41.47	437	337	0.92	-0.73	78.5	78.3
27	25.24	437	409	0.57	-1.72	93.6	93.6
28	47.53	437	292	0.85	-1.97	75.7	71.6
29	52.91	437	246	0.96	-0.63	69.1	68.6
30	33.96	437	379	0.65	-2.22	87.6	86.8
31	53.36	437	242	0.83	-3.18	73.9	68.5
32	57.84	437	202	1.2	3.38	61.3	68.4
33	34.61	437	376	0.77	-1.42	87	86.1
34	39.1	437	352	0.89	-0.82	83.1	81.1
35	48.02	437	288	0.99	-0.11	73	71.2
36	45.24	437	310	0.97	-0.34	75.1	73.9
37	65.38	437	138	1.15	1.86	70.7	73.1
38	75.94	437	69	1.7	3.89	83.1	84.6
39	48.38	437	285	1.03	0.41	73.2	70.9
40	48.86	437	281	0.85	-2.23	74.1	70.6
41	67.75	437	120	1.33	3.31	73.2	75.6
42	69.43	437	108	1.46	4.01	75.1	77.4
43	74.07	437	79	1.89	5.3	79.4	82.6
44	55.6	437	222	1.06	1.11	65.9	68.2
45	59.19	437	190	1.03	0.53	68.9	68.8
46	42.21	437	332	0.8	-1.97	80.5	77.4
47	55.38	437	224	1.3	5.06	58.6	68.2
48	58.96	437	192	0.87	-2.32	73.5	68.7
49	67.61	437	121	1.61	5.71	69.8	75.4
50	58.51	437	196	1.08	1.36	65.9	68.6

Table 26.
Three Types of Problematic Items.

		Items with Incorrect Answer Keys	Items Should Be Deleted or Revised	Items Should Be Kept in spite of Noticeable Off-variable Noises
Grade 1	F	None	None	Q1, Q5, Q10, Q16, Q18
	W	None	Q29, Q30	Q2, Q3, Q26, Q27, Q28
	S	None	Q26, Q30	Q25
Grade 2	F	None	None	Q16
	W*	None	Q27	None
	S	None	None	None
Grade 3	F	None	None	Q21
	W	None	None	Q32, Q48
	S	Q 22, Q40	None	Q32, Q48
Grade 4	F	Q 21, Q32	None	Q35
	W	None	None	Q32, Q48
	S	None	None	None
Grade 5	F	None	None	None
	W	None	None	Q26
	S	None	Q13, Q41	None
Grade 6	F	Q32	None	Q28
	W	None	None	None
	S	None	None	Q25
Grade 7	F	None	None	None
	W	None	None	None
	S	Q18	None	Q22, Q36, Q37, Q48
Grade 8	F	None	Q38	Q36
	W	None	Q15	None
	S	None	None	Q43, Q49

Additional notes for Table 26:

1. First Grade Fall Measure: These five problematic items are considered content appropriate for first grade students. The unexpected patterns are due to missing data. One of the students answered only three of the 20 questions and got them all correctly. WINSTEPS estimated this student's skill level as equivalent as students who answered 18 or 19 students correctly. This student did not answer Questions 1, 5 10, and 16. As a result, the average skill levels of students who answered Questions 1, 5, 10 and 16 correctly is lower than the one and only student who did not answer the questions. This might not be questions for other test items because more than one student missed the questions.
2. The Eighth Grade Fall 2006 has two problematic items. Both items are among the most difficult items in this measure and yet the answer keys are correct even though students who

answered correctly had lower average skill level than students who chose the most popular options. The author suggested keeping the items as they are.

3. First Grade Winter 2007: All five application problems have noticeable off-variables noises interfering with useful information. One of the sources for off-variable noise might come from the students' reading ability. Question 29 is problematic because they can be more than one justifiable answer. The problem did not specify the pattern shown is complete cycles of a pattern. The pattern "123____ 23" can be embedded in more than one possible pattern. The following are two of the examples: "123423453456" and "123223323423". The author recommended to revise the question as followed: "What is missing in this pattern? 1, 2, 3, 1, 2, 3, 1, 2, ___, 1, 2, 3". Question 30 might be too easy for first grade students.
4. In the Grade 2 Winter Measure, the author found two students who actually took the first grade test. The assertion was based on the students' response patterns that had 25 items dichotomously scored items and five multiple-choice items, while the second grade test had 50 multiple-choice items. The author deleted these two sets of data because it is not valid to include them into the data analysis. The author recommends that the districts include these students taking off-grade tests in the data set based on the test they take, not the grade that they are in.
5. Grade 2 Winter 2007 measure: Q27 asked what would be a reasonable to measure a weight of a car. Both the kilograms and tons can be justifiable answers. The answer provided is kilograms. The author recommends changing the distractor from ton to gram or milligram.
6. Grade 8 Winter 2007 measure: Q15 has the value of outfit mean square exceeding 2.00. The author recommends further review. The question asked students to simplify $(3m+1)(2m-4)$. One of the popular distractors is $5m-3$. Students who chose this distractor might mistake the multiplication problem to an addition problem. The other popular distractor is $6m^2-4$. Students who chose this distractor might mislead by the direction on the question to simplify. The students who answered the questions correctly have lower average skill levels than the students who chose these two popular distractors. There were only 40 out of 456 participating students choosing the correct answers. The author recommends to revise the question as followed:

Find the product of $(3m+1) \times (2m-4)$.

7. First Grade Spring 2007 measure: Question 1 provides overly predictive information. It could mislead the consumer to over-estimate the quality of that test item. Questions 26 was a relatively easy item, 94% of students who answered it correctly. Small number of students chose the incorrect answer and the variance of their ability level was large. The author suggested replacing the item. Q30 asked students to the find the greatest difference of improvements among 4 students over 4 years time. The data presented in the 4 by 4 format, it is a very difficult item for the first graders. I recommended simplifying the question, changing the format in 4 by 2 format (4 students and 2 years).
8. Second Grade Spring 2007 measure: Question 21 has the value of outfit mean squares exceeding 2.0. The question asks “What shape is NOT in the picture?” The students can make mistakes by not paying attention to the word NOT. Students who chose incorrect answer (rectangle) have higher average skill level than the students who chose the correct answer (circle). This question requires students decoding advanced vocabulary words from the domain of geometry. If students cannot decode the words: circle, triangle, square and rectangle, they cannot answer the question correctly. The author suggests including a pictorial representation next to the words circle, triangle, square and rectangle.
9. The author highlighted the several questions in yellow highlight. These questions shared the following characteristics: (a) the questions are aligned with grade-level standards on the Oregon Common Curriculum Goals. (b) The author compared the questions in the final version and the answers provided and determined that the answer keys were correct. (c) Students who answered correctly had lower average skill level than students who chose the most popular distractor. The author suggested further inspection to see if there was a discrepancy between the version of protocols being used.
10. Grade 5 Spring 2007 measures: Question 21 might be biased against certain groups of students because they don't know the words scooter and skateboard. The wording of Question 41 is confusing. The author recommended revising. Q37 had an incorrect answer key. After correction, Q37 is a constructive item.
11. Grade 6 Spring measure: Question 21 asked $\frac{1}{4} \times \frac{3}{7}$. The answer provided (3/28 – A) is the correct. However, the 295 students who choose 4/11 has the higher average skills level (58.75) than the 95 students who chose the correct answer (50.72). Question has an incorrect

answer (C). The correct answer should have been D. After correction, the item becomes a productive item.

12. Grade 7 Spring 2007 measures: Question 18 has incorrect key (E), it should be C. Question 21 is “ $5/6 \div \frac{1}{4} = \underline{\hspace{2cm}}$ ” . Question 22 might have off-variable noises because how the problem presented on the paper.

$$X+Y^*$$

$$Y-X$$

When $X=3$, $Y=5$.

Some students chose 16, because $(3+5)*(5-3) = 16$.

13. Grade 8 Spring 2007 measure: Question 21 is the same s Grade 7 Spring measure Q21.